10/531632
JC13 Rec'd PCT/PTO 14 APR 2005

Our Docket No.: 51876P839
Express Mail No.: EV372846822US

UTILITY APPLICATION FOR UNITED STATES PATENT

FOR

METHOD FOR GENERATING AND CONSUMING 3-D AUDIO SCENE WITH EXTENDED SPATIALITY OF SOUND SOURCE

Inventor(s):

Jeong Il SEO
Dae Young JANG
Kyeong Ok KANG
Jin-Woong KIM
Chieteuk AHN

Blakely, Sokoloff, Taylor & Zafman LLP
12400 Wilshire Boulevard, 7th Floor
Los Angeles, CA 90025
Telephone: (310) 207-3800

4\RRTS

10/531632

JC13 Rec'd PCT/PTO 14 APR 2005

METHOD FOR GENERATING AND CONSUMING 3-D AUDIO SCENE WITH EXTENDED SPATIALITY OF SOUND SOURCE

Description

Technical Field

The present invention relates to a method for generating and consuming a three-dimensional audio scene having sound source whose spatiality is extended; and, more particularly, to a method for generating and consuming a three-dimensional audio scene to extend the spatiality of sound source in a three-dimensional audio scene.

Background Art

Generally, a content providing server encodes contents in a predetermined encoding method and transmits the encoded contents to content consuming terminals that consume the contents. The content consuming terminals decode the contents in a predetermined decoding method and output the transmitted contents.

Accordingly, the content providing server includes an encoding unit for encoding the contents and a transmission unit for transmitting the encoded contents. On the other hand, the content consuming terminals includes a reception unit for receiving the transmitted encoded contents, a decoding unit for decoding the encoded contents, and an output unit for outputting the decoded contents to users.

Many encoding/decoding methods of audio/video signals are known so far. Among them, an encoding/decoding method based on Moving Picture Experts Group 4 (MPEG-4) is widely used these days. MPEG-4 is a technical standard for data compression and restoration technology defined by the MPEG to transmit moving pictures at a low transmission rate.

According to MPEG-4, an object of an arbitrary shape

can be encoded and the content consuming terminals consume a scene composed of a plurality of objects. Therefore, MPEG-4 defines Audio Binary Format for Scene (Audio BIFS) with a scene description language for designating a sound object expression method and the characteristics thereof.

Meanwhile, along with the development in video, users want to consume contents of more lifelike sounds and video quality. In the MPEG-4 AudioBIFS, an AudioFX node and a DirectiveSound node are used to express spatiality of a three-dimensional audio scene. In these nodes, modeling of sound source is usually depended on point-source. Point-source can be described and embodied in a three-dimensional sound space easily.

Actual point-sources, however, tend to have a dimension more than two, rather than to be a point of literal meaning. More important thing here is that the shape of the sound source can be recognized by human beings, which is disclosed by J. Baluert, "Spatial Hearing," the MIT Press, Cambridge Mass, 1996.

For example, a sound of waves dashing against the coastline stretched in a straight line can be recognized as a linear sound source instead of a point sound source. To improve the sense of the real of the three-dimensional audio scene by using the AudioBIFS, the size and shape of the sound source should be expressed. Otherwise, the sense of the real of a sound object in the three-dimensional audio scene would be damaged seriously.

That is, the spatiality of a sound source could be described to endow a three-dimensional audio scene with a sound source which is of more than one-dimensional.

Disclosure of Invention

It is, therefore, an object of the present invention to provide a method for generating and consuming a three-

dimensional audio scene having a sound source whose spatiality is extended by adding sound source characteristics information having information on extending the spatiality of the sound source to three-dimensional audio scene description information.

The other objects and advantages of the present invention can be easily recognized by those of ordinary skill in the art from the drawings, detailed description and claims of the present specification.

In accordance with one aspect of the present invention, there is provided a method for generating a three-dimensional audio scene with a sound source whose spatiality is extended, including the steps of: a) generating a sound object; and b) generating three-dimensional audio scene description information including sound source characteristics information for the sound object, wherein the sound source characteristics information includes spatiality extension information of the sound source which is information on the size and shape of the sound source expressed in a three-dimensional space.

In accordance with one aspect of the present invention, there is provided a method for consuming a three-dimensional audio scene with a sound source whose spatiality is extended, including the steps of: a) receiving a sound object and three-dimensional audio scene description information including sound source characteristics information for the sound object; and b) outputting the sound object based on the three-dimensional audio scene description information, wherein the sound source characteristics information includes spatiality extension information which is information on the size and shape of a sound source expressed in a three-dimensional space.

Brief Description of Drawings

The above and other objects and features of the present invention will become apparent from the following description of the preferred embodiments given in conjunction with the accompanying drawings, in which:

Fig. 1 is a diagram illustrating various shapes of sound sources;

Fig. 2 is a diagram describing a method for expressing spatial sound source by grouping successive point sound sources;

Fig. 3 shows an example where spatiality extension information is added to a "DirectiveSound" node of AudioBIFS in accordance with the present invention;

Fig. 4 is a diagram illustrating how a sound source is extended in accordance with the present invention; and

Fig. 5 is a diagram depicting the distributions of point sound sources based on the shapes of various sound sources in accordance with the present invention.

Best Mode for Carrying Out the Invention

Other objects and aspects of the invention will become apparent from the following description of the embodiments with reference to the accompanying drawings, which is set forth hereinafter.

Following description exemplifies only the principles of the present invention. Even if they are not described or illustrated clearly in the present specification, one of ordinary skill in the art can embody the principles of the present invention and invent various apparatuses within the concept and scope of the present invention.

The use of the conditional terms and embodiments presented in the present specification are intended only to make the concept of the present invention understood, and

they are not limited to the embodiments and conditions mentioned in the specification.

In addition, all the detailed description on the principles, viewpoints and embodiments and particular embodiments of the present invention should be understood to include structural and functional equivalents to them. The equivalents include not only currently known equivalents but also those to be developed in future, that is, all devices invented to perform the same function, regardless of their structures.

For example, block diagrams of the present invention should be understood to show a conceptual viewpoint of an exemplary circuit that embodies the principles of the present invention. Similarly, all the flowcharts, state conversion diagrams, pseudo codes and the like can be expressed substantially in a computer-readable media, and whether or not a computer or a processor is described distinctively, they should be understood to express various processes operated by a computer or a processor.

Functions of various devices illustrated in the drawings including a functional block expressed as a processor or a similar concept can be provided not only by using hardware dedicated to the functions, but also by using hardware capable of running proper software for the functions. When a function is provided by a processor, the function may be provided by a single dedicated processor, single shared processor, or a plurality of individual processors, part of which can be shared.

The apparent use of a term, 'processor', 'control' or similar concept, should not be understood to exclusively refer to a piece of hardware capable of running software, but should be understood to include a digital signal processor (DSP), hardware, and ROM, RAM and non-volatile memory for storing software, implicatively. Other known and commonly used hardware may be included therein, too.

In the claims of the present specification, an element expressed as a means for performing a function described in the detailed description is intended to include all methods for performing the function including all formats of software, such as combinations of circuits for performing the intended function, firmware/microcode and the like. To perform the intended function, the element is cooperated with a proper circuit for performing the software. The present invention defined by claims includes diverse means for performing particular functions, and the means are connected with each other in a method requested in the claims. Therefore, any means that can provide the function should be understood to be an equivalent to what is figured out from the present specification.

Other objects and aspects of the invention will become apparent from the following description of the embodiments with reference to the accompanying drawings, which is set forth hereinafter. The same reference numeral is given to the same element, although the element appears in different drawings. In addition, if further detailed description on the related prior arts is determined to blur the point of the present invention, the description is omitted. Hereafter, preferred embodiments of the present invention will be described in detail.

Fig. 1 is a diagram illustrating various shapes of sound sources. Referring to Fig. 1, a sound source can be a point, a line, a surface and space having a volume. Since sound source has an arbitrary shape and size, it is very complicated to describe the sound source. However, if the shape of the sound source to be modeled is controlled, the sound source can be described less complicatedly.

In the present invention, it is assumed that point sound sources are distributed uniformly in the dimension of a virtual sound source in order to model sound sources of various shapes and sizes. As a result, the sound sources

of various shapes and sizes can be expressed as continuous arrays of point sound sources. Here, the location of each point sound source in a virtual object can be calculated using a vector location of a sound source which is defined in a three-dimensional scene.

When a spatial sound source is modeled with a plurality of point sound sources, the spatial sound source should be described using a node defined in AudioBIFS. When the node defined in AudioBIFS, which will be referred to as an AudioBIFS node, is used, any effect can be included in the three-dimensional scene. Therefore, an effect corresponding to the spatial sound source can be programmed through the AudioBIFS node and inserted to the three-dimensional scene.

However, this requires very complicated Digital Signal Processing (DSP) algorithm and it is very troublesome to control the dimension of the spatial sound source.

Also, the point sound sources distributed in a limited dimension of an object are grouped using the AudioBIFS, and the spatial location and direction of the sound sources can be changed by changing the sound source group. First of all, the characteristics of the point sound sources are described using a plurality of "DirectiveSound" node. The locations of the point sound sources are calculated to be distributed on the surface of the object uniformly.

Subsequently, the point sound sources are located with a spatial distance that can eliminate spatial aliasing, which is disclosed by A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," J. Aoust. Soc. Am., Vol. 93, No. 5 on pages from 2764 to 2778, May, 1993. The spatial sound source can be vectorized by using a group node and grouping the point sound sources.

Fig. 2 is a diagram describing a method for expressing spatial sound source by grouping successive point sound

sources. In the drawing, a virtual successive linear sound source is modeled by using three point sound sources which are distributed uniformly along the axis of the linear sound source.

The locations of the point sound sources are determined to be (x_0-dx, y_0-dy, z_0-dz), (x_0, y_0, z_0), and (x_0+dx, y_0+dy, z_0+dz) according to the concept of the virtual sound source. Here, dx, dy and dz can be calculated from a vector between a listener and the location of the sound source and the angle between the direction vectors of the sound source, the vector and the angle which are defined in an angle field and a direction field.

Fig. 2 describes a spatial sound source by using a plurality of point sound sources. AudioBIFS appears it can support the description of a particular scene. However, this method requires too much unnecessary sound object definition. This is because many objects should be defined to model one single object.

When it is told that the genuine object of hybrid description of Moving Picture Experts Group 4 (MPEG-4)is more object-oriented representations, it is desirable to combine the point sound sources, which are used for model one spatial sound source, and reproduce one single object.

In accordance with the present invention, a new field is added to a "DirectiveSound" node of the AudioBIFS to describe the shape and size attributes of a sound source. Fig. 3 shows an example where spatiality extension information is added to a "DirectiveSound" node of AudioBIFS in accordance with the present invention.

Referring to Fig. 3, a new rendering design corresponding to a value of a "SourceDimensions" field is applied to the "DirectiveSound" node. The "SourceDimensions" field also includes shape information of the sound source. If the value of the "SourceDimensions"

field is "0,0,0", the sound source becomes one point, no additional technology for extending the sound source is applied to the "DirectiveSound" node. If the value of the "SourceDimensions" field is a value other than "0,0,0", the dimension of the sound source is extended virtually.

The location and direction of the sound source are defined in a location field and a direction field, respectively, in the "DirectiveSound" node. The dimension of the sound source is extended in vertical to a vector defined in the direction field based on the value of the "SourceDimensions" field.

The "location" field defines the geometrical center of the extended sound source, whereas the "SourceDimensions" field defines the three-dimensional size of the sound source. In short, the size of the sound source extended spatially is determined according to the values of Δx, Δy and Δz.

Fig. 4 is a diagram illustrating how a sound source is extended in accordance with the present invention. As illustrated in the drawing, the value of the "SourceDimensions" field is $(0, \Delta y, \Delta z)$, Δy and Δz being not zero ($\Delta y \neq 0$, $\Delta z \neq 0$). This indicates a surface sound source having an area of $\Delta y \times \Delta z$.

The illustrated sound source is extended in a direction vertical to a vector defined in the "direction" field based on the values of the "SourceDimensions" field, i.e., $(0, \Delta y, \Delta z)$, and thereby forming a surface sound source. As shown in the above, when the dimension and location of a sound source is defined, the point sound sources are located on the surfaces of the extended sound source. In the present invention, the locations of the point sound sources are calculated to be distributed on the surfaces of the extended sound source uniformly.

Figs. 5A to 5C are diagrams depicting the

distributions of point sound sources based on the shapes of various sound sources in accordance with the present invention. The dimension and distance of a sound source are free variables. So, the size of the sound source that can be recognized by a user can be formed freely.

For example, multi-track audio signals that are recorded by using an array of microphones can be expressed by extending point sound sources linearly as shown in Fig. 5A. In this case, the value of the "SourceDimensions" field is $(0, 0, \Delta z)$.

Also, different sound signals can be expressed as an extension of a point sound source to generate a spread sound source. Figs. 5B and 5C show a surface sound source expressed through the spread of the point sound source and a spatial sound source having a volume. In case of Fig. 5B, the value of the "SourceDimensions" field is $(0, \Delta y, \Delta z)$ and, in case of Fig. 5C, the value of the "SourceDimensions" field is $(\Delta x, \Delta y, \Delta z)$.

As the dimension of a spatial sound source is defined as described in the above, the number of the point sound sources (i.e., the number of input audio channels) determines the density of the point sound sources in the extended sound source.

If an "AudioSource" node is defined in a "source" field, the value of a "numChan" field may indicate the number of used point sound sources. The directivity defined in "angle," "directivity" and "frequency" fields of the "DirectiveSound" node can be applied to all point sound sources included in the extended sound source uniformly.

The apparatus and method of the present invention can produce more effective three-dimensional sounds by extending the spatiality of sound sources of contents.

While the present invention has been described with respect to certain preferred embodiments, it will be

apparent to those skilled in the art that various changes and modifications may be made without departing from the scope of the invention as defined in the following claims.